

### Datenbanken und Informationsverwaltungssysteme: Probleme ihrer Implementierung auf nur einmal beschreibbaren Speichermedien

Gathmann, Immo

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

#### Empfohlene Zitierung / Suggested Citation:

Gathmann, I. (1987). Datenbanken und Informationsverwaltungssysteme: Probleme ihrer Implementierung auf nur einmal beschreibbaren Speichermedien. *Historical Social Research*, 12(1), 76-87. <https://doi.org/10.12759/hsr.12.1987.1.76-87>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see: <https://creativecommons.org/licenses/by/4.0>

## software

---

### DATENBANKEN UND INFORMATIONSVERWALTUNGSSYSTEME PROBLEME IHRER IMPLEMENTIERUNG AUF NUR EINMAL BESCHREIBBAREN SPEICHERMEDIEN

I. Gathmann (\*)

*Abstract: Die folgenden Kapitel befassen sich mit dem neuesten Vertreter der Hintergrundspeicher, der numerischen optischen Platte (NOP). Kapitel 3 erklärt den Aufbau und die Funktionsweise dieses Speichermediums und beschreibt dessen Vor- und Nachteile im Vergleich zur Magnetplatte. In Kapitel 4 wird ein Ausblick über die Struktur zukünftiger Datenhaltungssysteme und die Einbindung der NOP in ein solches System gegeben.*

#### I. Unterschiede zwischen Datenbanksystemen (DBS) und Informationsverwaltungssystemen (IRS)

Ein Datenhaltungssystem kann nach ANSI/SPARC (1975) wie in Bild 1.1. dargestellt aufgebaut werden.

Das **externe** Modell gibt die Sicht auf Teile der Umwelt wieder, wie sie vom Benutzer benötigt werden. Das **konzeptuelle** Modell beschreibt die Gesamtheit der zu verwaltenden Daten. Das **Interne** Modell legt fest, wie die Daten im Speicher abzulegen sind.

Die Transformationsregeln verbinden die beschriebenen Schichten, indem sie festlegen, wie ein Objekt eines Modells von Objekten des tieferliegenden Modells aufgebaut wird (Schlageter/Stucky 1983).

DBMS: Das Datenbankverwaltungssystem sorgt neben der Verwaltung der Daten für die Synchronisation mehrerer Benutzer, sowie für die Integrität und den Schutz der Daten.

BS: Betriebssystem

Betrachtet man, neben dem in Bild 1.1. dargestellten Schichtenmodell für Datenhaltungssysteme die in Bild 1.2 dargestellten globalen DV-Schichten (Wiese 1985), so kann man die in Bild 1.3 gezeigten Zuordnungen treffen.

Wie Bild 1.3 entnommen werden kann, belegt ein DBS mehr den Bereich der systemnahen Software.

Dies liegt zum einen daran, daß DBS so ausgelegt sind, daß Änderungen und die damit zusammenhängenden Komponenten Pufferspeicher- und Transaktionsverwaltung schneller vollzogen werden können (eine Transaktion ist eine Folge von Anweisungen,

---

(\*) Address all communications to: I. Gathmann, Büro für Archäologie und Informatik in den Kulturwissenschaften, D-7500 Karlsruhe 1.

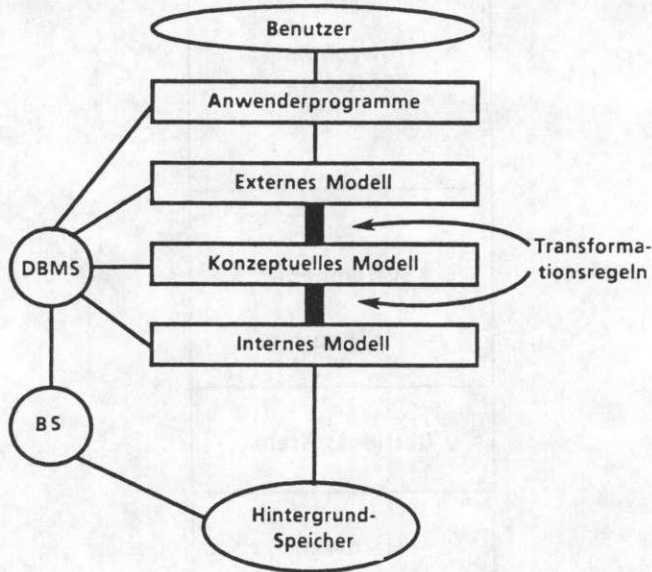


Bild 1.1

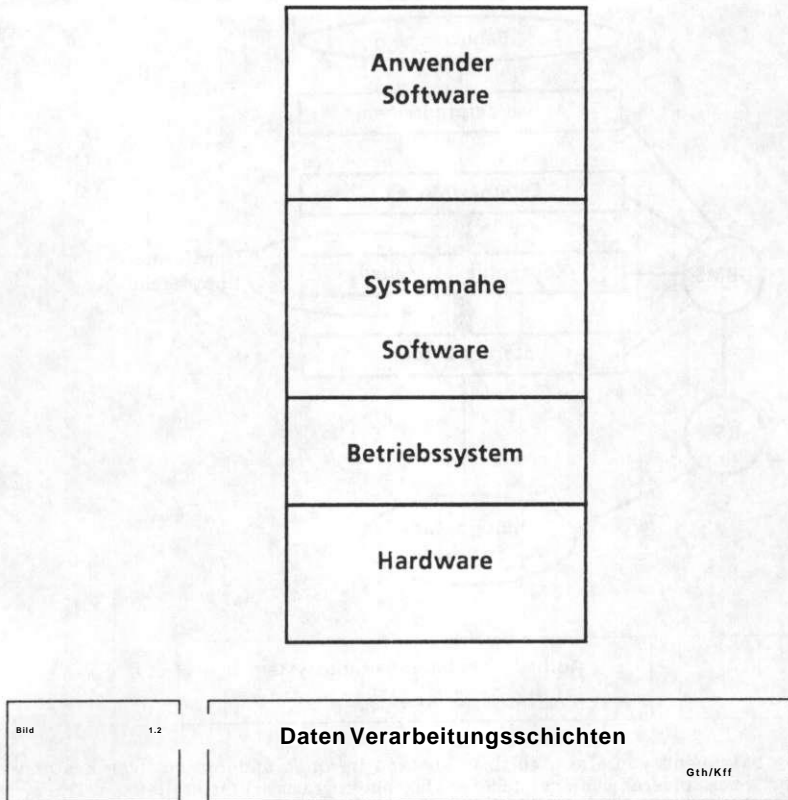
### Aufbau eines Datenhaltungssystem nach ANSI

GfH 11

die eine Datenbank von einem gültigen Zustand in einen anderen gültigen Zustand überführt); zum anderen bedienen DBS vorwiegend Programmschnittstellen und Transaktionen mit entsprechenden Sicherungsfunktionen. Letztendlich sind in DBS die Daten in strukturierter Form, die auch die Bedeutung der Daten bestimmt, abgespeichert (z.B. in indexsequentiellen Dateien), was eine eindeutige Anforderung eines Datums zuläßt.

In einem IRS liegen dagegen vorwiegend unformatierte Daten (Text) vor, denen bestimmte Suchbegriffe zugeordnet sind, deren Bedeutung in dem Dokument jedoch nicht immer eindeutig ist. So könnte z.B. das Stichwort *Tafel* in einem archäologischen Dokument sowohl auf ein seltenes Ausgrabungsstück als auch auf eine in dem Dokument vorgestellte Zeittafel hinweisen. Diese Mehrdeutigkeit macht die Wiedergewinnung von Daten in einem IRS aufwendiger als in einem DBS, denn um zu einem eindeutigen Ergebnis zu kommen, müssen weitere Daten (sogenannte Sekundärdaten) bereitgestellt werden. Ein IRS ist deshalb dialogorientiert und liegt mit seinen Funktionen – Ermittlung von Deskriptoren und deren Beziehungen untereinander, Bereitstellung von Blätterfunktionen in Wörterbüchern (Thesauri) und Abfragesprache – mehr in der obersten DV-Schicht.

Die Entwicklung geht jedoch dahin, die jeweils typischen Eigenschaften eines DBS und eines IRS zu vereinen. Für ein archäologisches Datenhaltungssystem wäre z.B. die Ansprache von Dokumenten bezüglich ihres Inhaltes (ein Ausstellungsstück betreffend) genauso von Interesse wie die reine Verwaltung von Inventardaten.



## II. Das periphere Speichermedium

Ein wesentlicher Bestandteil eines Datenhaltungssystems ist das zugehörige periphere Speichermedium, der sogenannte Sekundärspeicher. Er trägt entscheidend zur Effizienz des Gesamtsystems bei. Die leistungsentscheidenden Parameter sind dabei die Aufzeichnungsdichte, die Transferdatenblockgröße, die Zugriffszeit und die Datensicherheit (Schweppe 1985).

Die Aufzeichnungsdichte Bit/Inch bestimmt die Kapazität des Speichermediums sowie dessen Aufbau. So können heute auf einer 5 1/4 Zoll Magnetplatte ebensoviel Daten gespeichert werden wie vor 10 Jahren auf einer 12 Zoll Magnetplatte. Die Aufzeichnungsdichte liegt bei Magnetplatten, den zur Zeit dominierenden Speichermedien, zwischen 7000 und 14000 Bit/Inch, womit eine Speicherkapazität von bis zu 300 MByte erreicht werden kann.

Die Größe eines Datenblocks zum Transfer von Daten vom Hauptspeicher des Rechners zum Sekundärspeicher und umgekehrt bestimmt den internen Aufwand, der beim Zugriff auf das externe Speichermedium betrieben werden muß. Zwar werden bei

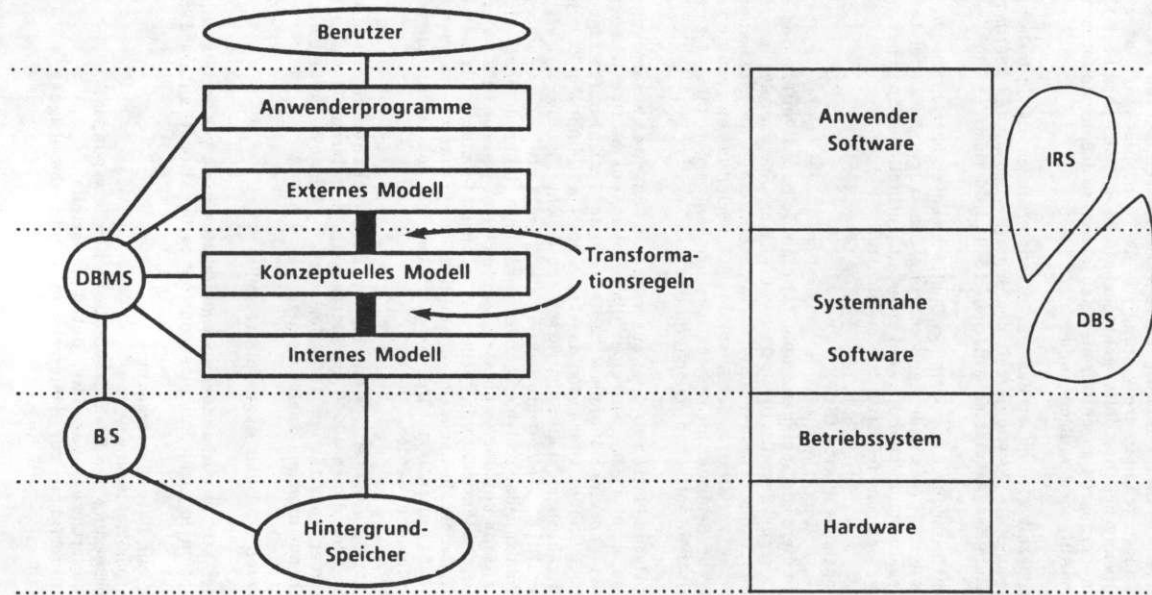


Bild 1.3

Zuordnung Datenbankschichten

Datenverarbeitungsschichten

Gth.k.H

großen Datenblöcken (1 - 8 KByte) auch mehr irrelevante Daten gelesen als bei kleineren Datenblöcken, deren genaue Adressierung z.B. über eine Indexdatei aber einen höheren internen Aufwand bedeutet. Man geht deshalb dazu über, periphere Koprozessoren, die den Speicherzugriff durchführen, einzusetzen, so daß der Hauptprozessor sich derweil anderen Aufgaben widmen kann.

Die mittlere Zugriffszeit wird neben dem Datentransfer durch die Umdrehungszeit des Speichermediums und die Positionierzeit des Schreib-/Lesekopfes bestimmt. Sie beträgt bei Hochleistungsmagnetplatten zwischen 15 und 30 msec.

Die Datensicherheit kann durch das periphere Speichermedium bestenfalls nur gegen mechanische Fehler wie z.B. Erschütterung, Berührung, Staub usw. gewährleistet werden. Eine robuste Kapselung und/oder die Wahl eines unempfindlichen Speicherverfahrens (wie z.B. bei der optischen Platte) bieten sich an.

### III. Die numerische optische Platte (NOP)

Die optische Platte dient als Speichermedium für digitale Daten. Aufbau und Funktionsweise seien hier kurz dargestellt. Die optisch-numerische Speicherplatte darf nicht mit analogen Bild- oder Compact-Disc-Platten verwechselt werden, die inzwischen auch als Informationsspeicher, vor allem in der Ausbildung und Schulung eingesetzt werden. Diese Medien können vom Benutzer nicht selbst beschrieben werden. Bild 3.1 zeigt den Aufbau einer numerischen optischen Speicherplatte.

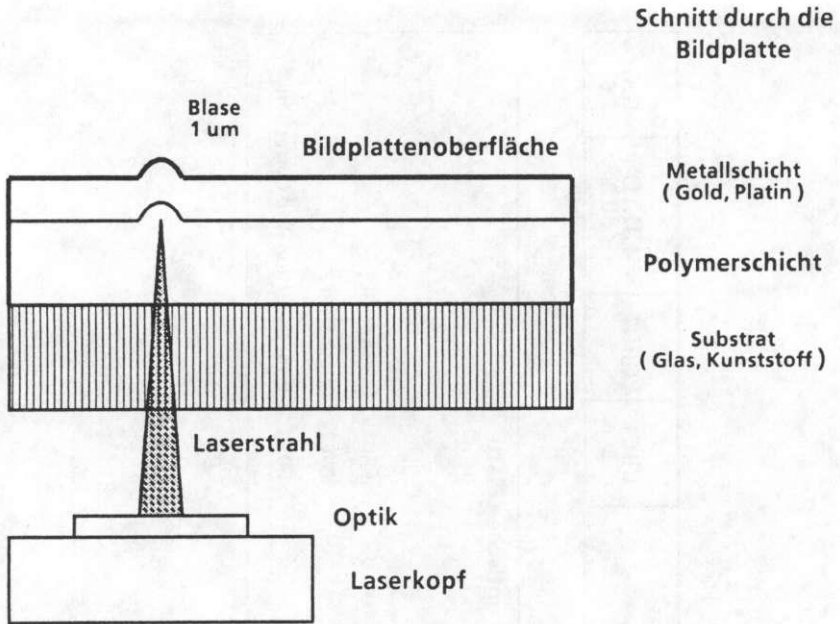
Die numerische optische Platte (NOP) ist aus mehreren Schichten zusammengesetzt ("Sandwich"-Bauweise, s. Bild 3.1). Auf einer Glas- oder Kunststoffschicht ist eine Polymerschicht gefolgt von einer dünnen Metallfolie (Gold, Platin) aufgetragen. Der Laserstrahl wird beim Schreiben oder Lesen durch die Optik auf einen Punkt von 1 µm gebündelt. Der Laserstrahl passiert die Substratschicht, durchdringt die Polymerschicht und trifft auf die Metallfolie. Beim Schreiben ist die Energie des Lasers so hoch (1,3 MWatt), daß die durchdrungene Polymerschicht verdampft. Dadurch beult sich die darüberliegende Metallfolie aus und es entsteht ein Bläschen von ca. 1 µm Durchmesser. Zum Lesen wird ein Laserstrahl von geringerer Intensität benutzt. Die Metallfolie reflektiert den Strahl, wobei, je nachdem, ob ein Bläschen vorhanden ist oder nicht, mehr oder weniger Licht reflektiert wird. Somit können "0" und "1" als elementare Informationen unterschieden werden.

Die NOP ist in Sektoren aufgeteilt, wie in Bild 3.2 dargestellt ist. 25 dieser Sektoren ergeben eine Spur, von denen 40000 pro Plattenseite existieren (Thomsen 1984). Dies ergibt eine Speicherkapazität von 1000 Megabyte = 1 Gigabyte je Seite der auswechselbaren optisch numerischen Speicherplatte.

#### 1. Gemeinsamkeiten von NOP und Magnetplatte (MP)

Um den Einsatz der NOP in vorhandenen Rechnersystemen zu gewährleisten, wurden bei ihrer Konzeption logische Einheiten der MP beibehalten (Fujitani 1984).

- Schreib-/Lesekopf zur Datenübertragung;
- Mechanik zur Kopfpositionierung;
- Elektronik: zur Umsetzung der Rechnerbefehle in Kopfbewegungen,  
zur Konvertierung paralleler Daten in serielle,  
Kontrollverfahren zur Fehlererkennung und -beseitigung;
- gleiche Schnittstelle.

**Bild 3.1**

### Arbeitsweise der numerische optischen Platte (NOP)

Gth/Kff

## 2. Unterschiede zwischen NOP und MP

Die gravierendsten Unterschiede betreffen die verschiedenen Verfahren zum Schreiben bzw. Lesen von Daten. Bei der MP erzeugt der Schreib-/Lesekopf ein dichtes Magnetfeld, um lokal die Richtung der magnetischen Einheiten zu ändern bzw. solche Veränderungen zu erkennen. Dazu ist es notwendig, daß der Kopf bis auf 0.1 um an die Platte herangeführt wird, da sonst die Streuung zu groß wäre. Bei der NOP dagegen steht der Kopf in einer Entfernung von 2 mm zur Platte. Dies hat zur Folge, daß die NOP wesentlich unempfindlicher gegen Staub und Erschütterung ist als die MP. Die NOP kann deshalb auch ohne Gefährdung der Daten leicht ausgewechselt werden, um z.B. eine Kopie anzufertigen; außerdem ist eine höhere Datensicherheit als bei der MP gewährleistet. Die unterschiedliche Art der Aufzeichnung bedingt auch eine um den Faktor 10 - 20 höhere Aufzeichnungsdichte bei der NOP. Nachteile der Aufzeichnungsart der NOP sind die bisherige Nichtlösbarkeit der einmal aufgetragenen Daten sowie die um den Faktor 10 höhere Zugriffszeit der NOP im Vergleich zur MP (siehe auch Vergleichstabelle Bild 3.3).

|          |         |          |               |           |           |            |          |
|----------|---------|----------|---------------|-----------|-----------|------------|----------|
| WCR<br>3 | HD<br>1 | LSA<br>3 | Daten<br>1024 | CRC1<br>2 | CRC2<br>2 | EDAC<br>80 | GAP<br>5 |
|----------|---------|----------|---------------|-----------|-----------|------------|----------|

**Bild 3.2**

**Sektoren der Numerischen optischen Platte**

WCR : Write Check Field

HD : Header

LSA : Logical Sector Address

CRC1 : Check Field 1

CRC2 : Check Field 2

EDAC : Error Detection And Correction



| Eigenschaften                            | Lebens-<br>dauer   | Aufzeich-<br>nungsdichte                 | Zugriffszeit | Leistungs-<br>aufnahme | Kosten /<br>MByte | Abstand des<br>S/L-Kopfes |
|--|--------------------|--|--------------|------------------------|-------------------|---------------------------|
| Magnetplatte<br>(Harddisk)               | max. 5 Jahre       | $12.2 \cdot 10^6 /$<br>inch <sup>2</sup> | 15-30 ms     | 1000W                  | \$60              | 0.1 mm                    |
| Optische<br>Speicherplatte<br>(Gigadisc) | größer 10<br>Jahre | $210 \cdot 10^6 /$<br>inch <sup>2</sup>  | 200 - 300 ms | 250 W                  | \$8               | 2 mm                      |

Bild 3.3

Gegenüberstellung einiger Eigenschaften von magnetischen und optisch-numerischen Speicherplatten

Da bei der NOP nur eine vergleichsweise leichte Kunststoffplatte bewegt werden muß, ist die Leistungsaufnahme der NOP im Vergleich zur MP um den Faktor 4 - 5 geringer. Die Kosten pro Byte belaufen sich bei der NOP auf \$ 8 bei der MP auf \$ 60. Nach der Einführungszeit ist mit noch geringeren Preisen zu rechnen.

#### **IV. Probleme der Datenverwaltung auf der nur einmal beschreibaren NOP**

Die Nichtlöschbarkeit einmal auf die NOP geschriebener Daten erfordert neue Speicherstrategien (z.B. Mehrfachclusterbildung im Hinblick auf Datenhaltungssysteme oder Partitionierung) sowie angepaßte oder neue Betriebssysteme, die das Problem der Nichtlöschbarkeit mit berücksichtigen (Scheele 1985).

So ist z.B. das Führen verketteter Listen unter Beibehaltung der ständigen "Update"-Fähigkeit nicht möglich. Auch die Führung eines stets aktuellen Inhaltsverzeichnisses bereitet Schwierigkeiten, denn wenn man jede neue Datei sofort eintragen würde, müßte jeweils ein Sektor von ca. 1 KByte belegt werden, wovon maximal 25 Bytes benötigt würden. Eine Lösung dieses Problems wäre z.B. die Verkettung mehrerer Dateien, bis soviele Einträge gesammelt sind, daß ein Sektor voll belegt werden kann.

Eine weitere Möglichkeit wäre die Führung des Inhaltsverzeichnisses auf einem anderen vorhandenen Speichermedium, z.B. MP oder Floppy Disc, das erst auf die NOP abgespeichert wird, wenn diese nahezu voll belegt ist.

Soll das Inhaltsverzeichnis von Beginn an auf der NOP geführt werden, empfiehlt es sich, die Daten z.B. von außen nach innen abzuspeichern, das Inhaltsverzeichnis dagegen von innen nach außen, da von vorneherein nicht bekannt ist, wieviel Platz jeweils benötigt wird. Durch das Lesen des Inhaltsverzeichnisses von außen nach innen (im Gegensatz zum Schreiben), können die jeweils aktuellen Informationen besonders schnell erreicht werden. So ergibt sich auch eine Möglichkeit zum Korrigieren und Ergänzen von Informationen, da veraltete Daten auf Grund der Struktur des Inhaltsverzeichnisses im Normalbetrieb gar nicht mehr erreicht werden.

Da für jede Änderung auf der NOP ein völlig neuer Speicherbereich zur Verfügung gestellt werden muß, muß sich das Betriebssystem auf den ständig schrumpfenden Speicherplatz einstellen. Außerdem muß es in der Lage sein, eine auf zahlreiche Sektoren verteilte Datei effizient wieder zusammensetzen zu können. Es erscheint daher sinnvoll, in periodischen Abständen eine Sammlung der noch gültigen Daten durchzuführen und sie auf eine zweite Platte zu kopieren.

Eine weitere Forderung an ein - auf die NOP zugeschnittenes - Betriebssystem wäre die Fähigkeit, zwischen sich laufend ändernden und dauernd festzuhaltenden Daten zu unterscheiden.

#### **V. Ausblick**

Die NOP wird die MP, auch wenn sie einmal wieder löschbar sein sollte, woran zur Zeit intensiv gearbeitet wird, nicht ersetzen können, denn dazu sind die Zugriffszeiten zu hoch. Aber sie wird die MP als sogenannter tertiärer "ON-LINE"-Speicher (Massenspeicher mit direkten Zugriffsmöglichkeiten) wirkungsvoll unterstützen können. Sie wird damit auf lange Sicht die bisherige tertiären "OFF-LINE"-Speichermedien wie Magnetbänder ersetzen (Massenspeicher, die erst durch mechanische oder manuelle Tätigkeit auf Anforderung des Rechners bereitgestellt werden) (Fujitani 1984).

Als typisches Anwendungsgebiet bietet sich die NOP aufgrund ihrer hohen Datensicherheit als Archivspeichermedium an, auf das im Vergleich zu anderen Archivspeichern wie Magnetbänder oder -kassetten schnell zugegriffen werden kann.

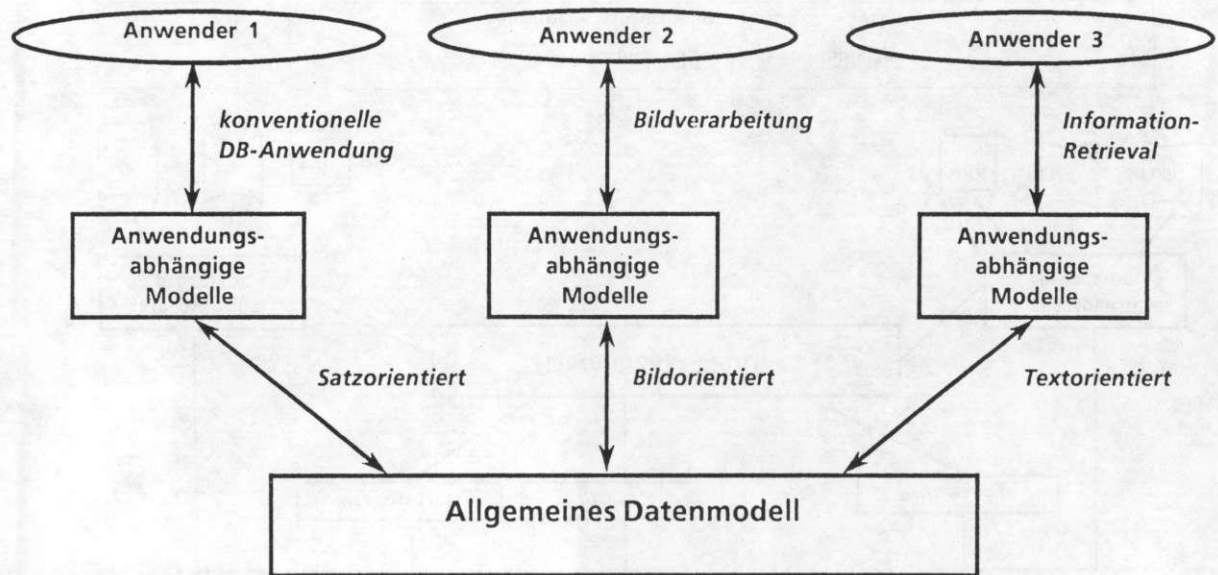


Bild 5.1

Systemarchitektur eines kombinierten  
Datenbank -- Informations-Retrieval -- Systems

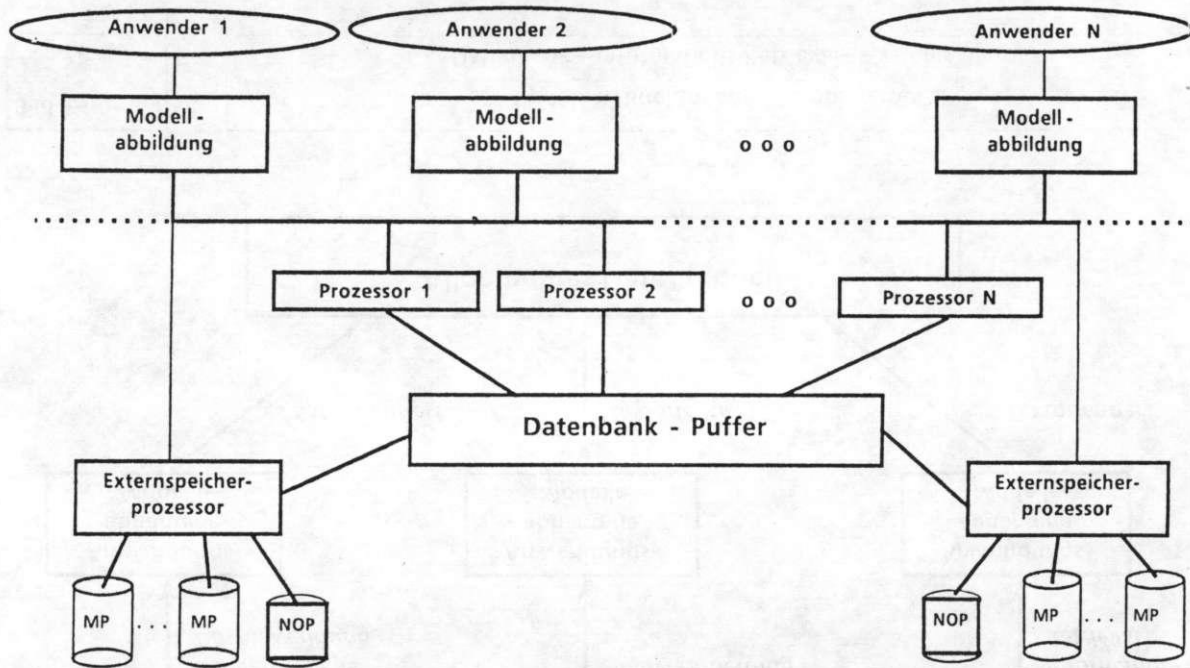


Bild 5.2

Prozessorarchitektur eines kombinierten  
Datenbank -- Informations-Retrieval -- Systems

Gith/KII

Ein weiteres mögliches Einsatzgebiet der NOP sind lokale Rechnernetze, in denen die NOP als Zentralspeicher fungieren könnte. Aber auch als Datenbasis verteilter IRS anstelle der zur Zeit noch zentralisierten IRS-Basen bietet sich die NOP an.

Wie schon in Kapitel 1 angedeutet, besteht heute die Tendenz, DBS und IRS zu vereinen. Es ergibt sich daher die Notwendigkeit der Entwicklung neuer Systemarchitekturen, in die auch die NOP als tertiärer "ON-LINE"-Speicher mit einbezogen wird. Wie ein solches System aussehen kann zeigt Bild 5.1 (nach Härder/Reuter 1985).

Dabei werden verschiedene Anwendungsprozesse auf ein entsprechendes Modell abgebildet. Diese mit sehr viel Semantik behafteten Modelle werden anschließend auf ein allgemeines Datenmodell, das in sich alle Speicherverwaltungs-, Darstellungs- und Zugriffstechniken vereint, abgebildet.

Um eine effiziente, zeitparallele Abarbeitung der verschiedenen Teiloperationen zu gewährleisten, ist eine entsprechend hohe Zahl aktiver Prozessoren notwendig. Als Externspeicher bietet sich neben den üblichen MP aufgrund ihrer extremen Speicherkapazität auch die NOP an, insbesondere für die Abspeicherung komplexer Strukturen wie z.B. von Bildern. Eine mögliche Prozessorarchitektur zeigt Bild 5.2 (nach Härder/Reuter 1985).

Im Vergleich zu einem konventionellen DBS ergeben sich folgende Unterschiede:

- Die einzelnen Datenmodelle sind mit mehr Semantik versehen;
- stärkere Trennung der logischen Modelle von den internen Darstellungstechniken;
- Parallelität bei Abarbeitung komplexer Operationen;
- mächtigere Suchverfahren, wie z.B. Suche aufgrund von Ähnlichkeiten;
- Einbindung von schnellen Archivspeichern mit hohen Kapazitäten;
- neue Speicherstrukturen.

## REFERENCES

- ANSI/SPARC Study Group on DBMS, Interim report 75-02-08. FDT (Bulletin of ACM-SIGMOD) 7 (1975), Nr. 2.
- Fujitani, L., Laser optical disk: The Coming revolution in on-line storage; in: CACM 27 (6), Juni 84, S. 546-554.
- Härder, T./A. Reuter, Architektur von Datenbanksystemen für Nonstandard-Anwendungen; Informatik Fachberichte 94, Springer Verlag 1985, S. 253-286.
- Schweppe, H., Hardwareunterstützung für Datenbanken in Büro, Technik und Wissenschaft; in: Informatik Fachbericht 94, Springer Verlag 1985, S. 287-308.
- Scheele, H.J, Elektronikpraxis Nr. 3, März 1985, S. 119-123.
- Schlageter, G./W. Stucky, Datenbanksysteme. Konzepte und Modelle, Teubner Studienbücher Informatik 1983.
- Thomson, Alcatel, Gigadisc, OEM Manual GD 1001, Oktober 1984.
- Wiese, L., Golem XT: Ein IRS aufbauend auf einem DBS; in: Informatik Fachberichte 94, Springer Verlag 1985, S. 217-233.